Interface and Model Operation

- Model requires Microsoft Office Professional '97
- Model can now be run in "batch" mode -- select a group of study areas and the model will cycle through them sequentially
- Computing requirements are similar to those for HM4
  - Fast Pentium processor recommended
  - 32MB of RAM minimum, 64 MB recommended

Counting, Locating, Clustering and Engineering Plant to Customers in the Hatfield Model, v5.0

PNR and Associates,
Jenkintown, PA

Hatfield Associates,
Boulder, CO

Sponsors: *AT&T*
MCI

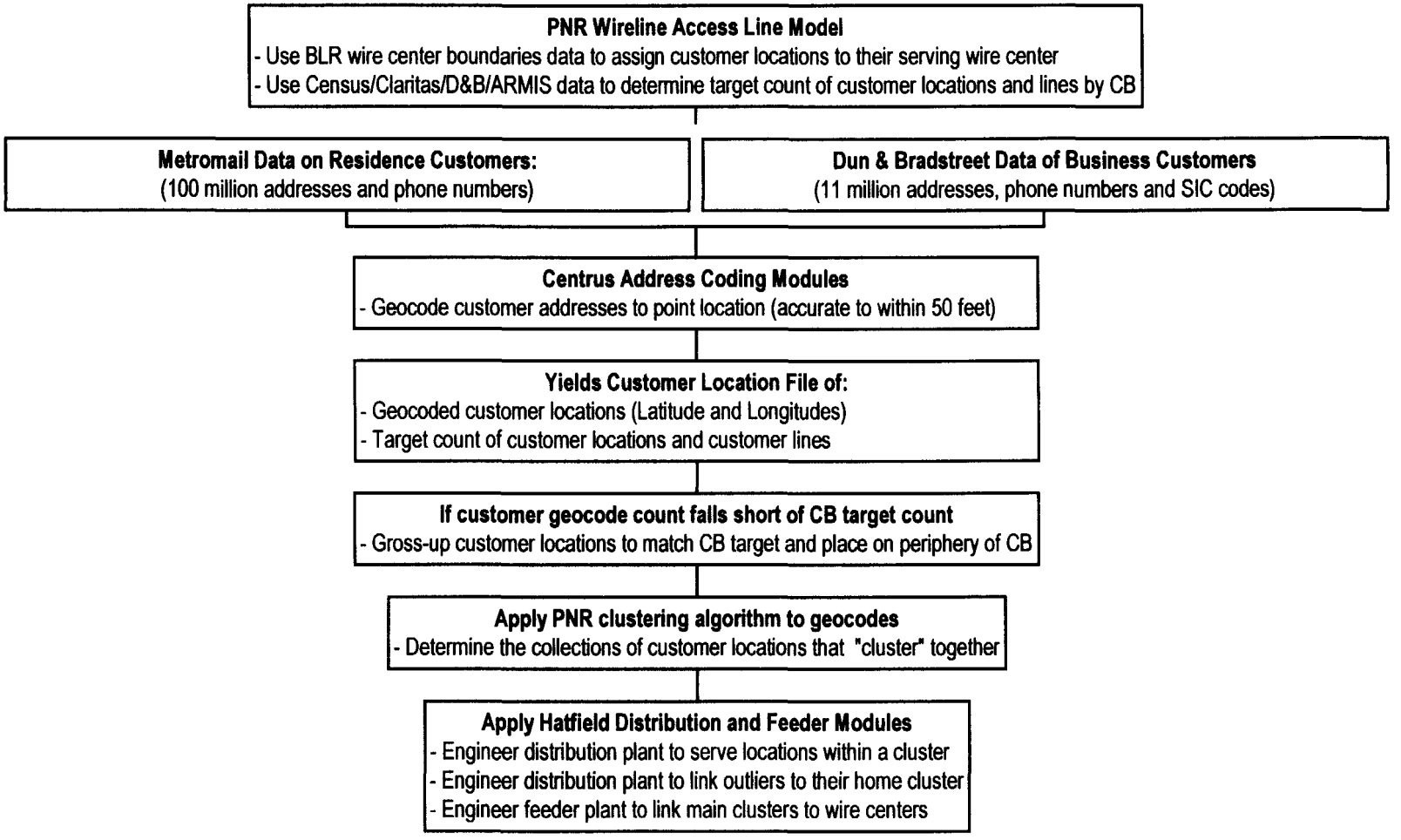
NARUC Convention
November 7-13, 1997

Process Stages

- Counting customers
  - PNR Wireline Access Line Model (WALM)
- Locating customers
  - Geocoding actual customer location
  - Location gross-up process
- Clustering process
  - Grouping customers that can be served in a single distribution area into distribution areas
- Engineering distribution plant
  - Placing cables to serve located customer clusters

Overview

Process Overview



PNR Wireline Access Line Model

- PNR Wireline Access Line Model (WALM) estimates number of:
  - Customer locations, and
  - Telephone lines

  by
  - Census Block (CB), and by
  - Wire Center (WC)

PNR WALM - Residence

- Determine a target number of residence locations by CB
  - Claritas 1996 household count projections at the Census Block Group (CBG) level are apportioned to the CB level based on 1990 U.S. Census distribution patterns
  - Consistency checks performed relative to Metromail customer location counts and county-level 1995 U.S. Census update data

PNR WALM - Residence (continued)

- Determine a target number of residence telephone lines by CB based on:
  - Residence customer location counts
  - Telephone penetration estimates based on statewide U.S. Census data as adjusted due to CBG-specific demographic factors such as age and income
  - "Second" line penetration based on age and income
- Target counts of households and lines are then normalized to match study area residence line totals as reported in:
  - ARMIS, NECA, USTA, RUS, etc.

PNR WALM - Business

- Determine a target number of business locations by CB based on:
  - Dun & Bradstreet (D&B) business establishment counts located to the CB level based on geocoded addresses
- Determine a target number of business telephone lines by CB based on:
  - Gross counts of business lines estimated from D&B business establishment counts, employee counts and SIC code information
  - Shares of lines that are single line business or multiline business based on employee counts and SIC codes

PNR WALM - Business (continued)

- Target counts of business locations, employees and lines are then normalized to match study area business line totals as reported in:
  - ARMIS, NECA, USTA, RUS, etc.

PNR WALM - WC Assignment

- BLR data on WC boundaries is primary basis for assignment of customer locations and lines to WCs
  - BLR boundaries are conformed to CB boundaries
  - CBs are then assigned to WCs based on these boundaries
  - Customer telephone number NPA-NXX and Bellcore LERG WC associations are used to verify BLR accuracy and resolve data inconsistencies
- All customer locations located in CBs that are associated with a WC are then assigned to that WC

Metromail Residential Data

- Large file of deliverable postal addresses and telephone numbers
  - Contains over 100 million household records
  - Covers over 90% all U.S. households
- Compiled from many information sources:
  - Telephone white pages, voter and automobile records, real estate transaction data, warranty cards, etc.
- Accuracy maintained
  - File is updated 65 times annually
  - Incorporates postal address standardization practices
  - National change of address processing

Dun & Bradstreet Business Data

- Information on more than 11 million business establishments (over 90% coverage)
  - Postal addresses
  - Telephone numbers
  - Counts of employees
  - Standard Industrial Classification (SIC) codes
- Compiled from many information sources:
  - Gov't. and trade orgs., financial institutions, etc.
- Accuracy maintained via:
  - Daily updates
  - Annual management interviews

Geocoding Process

- Centrus Address Coding Modules used to assign geocodes to business and residence customer addresses
- Process:
  - Standardize addresses to USPS specifications
  - Determine latitude and longitude coordinates (geocodes) of each address
  - Retain geocodes that are precise to six decimal places of accuracy (50 feet or less)
  - Discard less accurate geocodes

Geocoding Process (continued)

- Quality measures
  - Nationally, about 80% of all addresses successfully geocode to the six decimal place point level
  - Ungeocoded locations due to:
    - Faulty address records
    - Faulty road location (TIGER) files
    - P.O. Boxes, rural routes, etc.
- Geocoding success is improving rapidly due to:
  - Tightened postal requirements
  - Needs of non-postal shippers (e.g., UPS, Fedex)
  - E911 expansions

Geocoding Process (continued)

Note the difference in Census Ids. Those with A SX codes are 12 digit while ZXnX codes are only 10

STREET	CITY	ZIP	LONG	LAT	CENSUS ID	MATCH CODE	LOC CODE
42 HUMMINGBIRD LN	ALMONT	81210	-106.486424	38.781180	08051963600110	S80	**AS0**
213 RAINBOW RD	ALMONT	81210	-106.488218	38.788311	08051963600110	S80	**AS0**
954 COUNTY ROAD 744	ALMONT	81210	-106.767130	38.749458	08051963800124	S80	AS1
2809 COUNTY ROAD 38	GUNNISON	81230	-106.940758	38.470850	08051963600288	S80	**AS1**
18806 COUNTY ROAD 765	ALMONT	81210	-106.480000	38.723500	080519636001	S80	**ZB9E**
5866 COUNTY ROAD 62	GUNNISON	81230	-106.617400	38.439800	080519636001	S80	ZB9E
299 ZUNI TRL	GUNNISON	81230	-106.936300	38.504200	080519636002	S80	**ZC7K**
146 CACTUS HILL DR	GUNNISON	81230	-106.643400	38.467800	080519636001	S88	ZC7K
200 INDUSTRIAL PARK RD	GUNNISON	81230	-106.559900	38.504100	080519636001	S89	ZC7K
34 SLATE VIEW LN	CRESTED BUTTE	81224	-106.882500	38.807100	080519638001	S88	**ZT7J**
17300N N STATE HIGHWAY 1	ALMONT	81210	-106.846000	38.664600	080519636002	S82	**ZT9F**

Note the difference in Lat / Long assignments. Point geocodes go to the 6th decimal place

Address information is provided courtesy of Metromail, Inc.

Geocoding Process (continued)

- Location Codes
  - Location codes indicate the accuracy of the assigned geocode

Code	Description
AS0	Indicates a house range address geocode. This is the most accurate geocode available. *Best Location - offset 50'.*
AS1	Indicates a house range address geocode. Street side is unknown. *CB to the left of the intersection is assigned and location is in middle of street.*
AS3	Indicates address was interpolated onto a TIGER segment that did not originally contain address ranges. *CB to the left of intersection is assigned and location is in middle of street.*
ZXnX	Indicates location based on ZIP code data. This is NOT a *point* geocode and therefore is not used in Customer Location

Geocoding Process (continued)

Levels of geocoding in our example

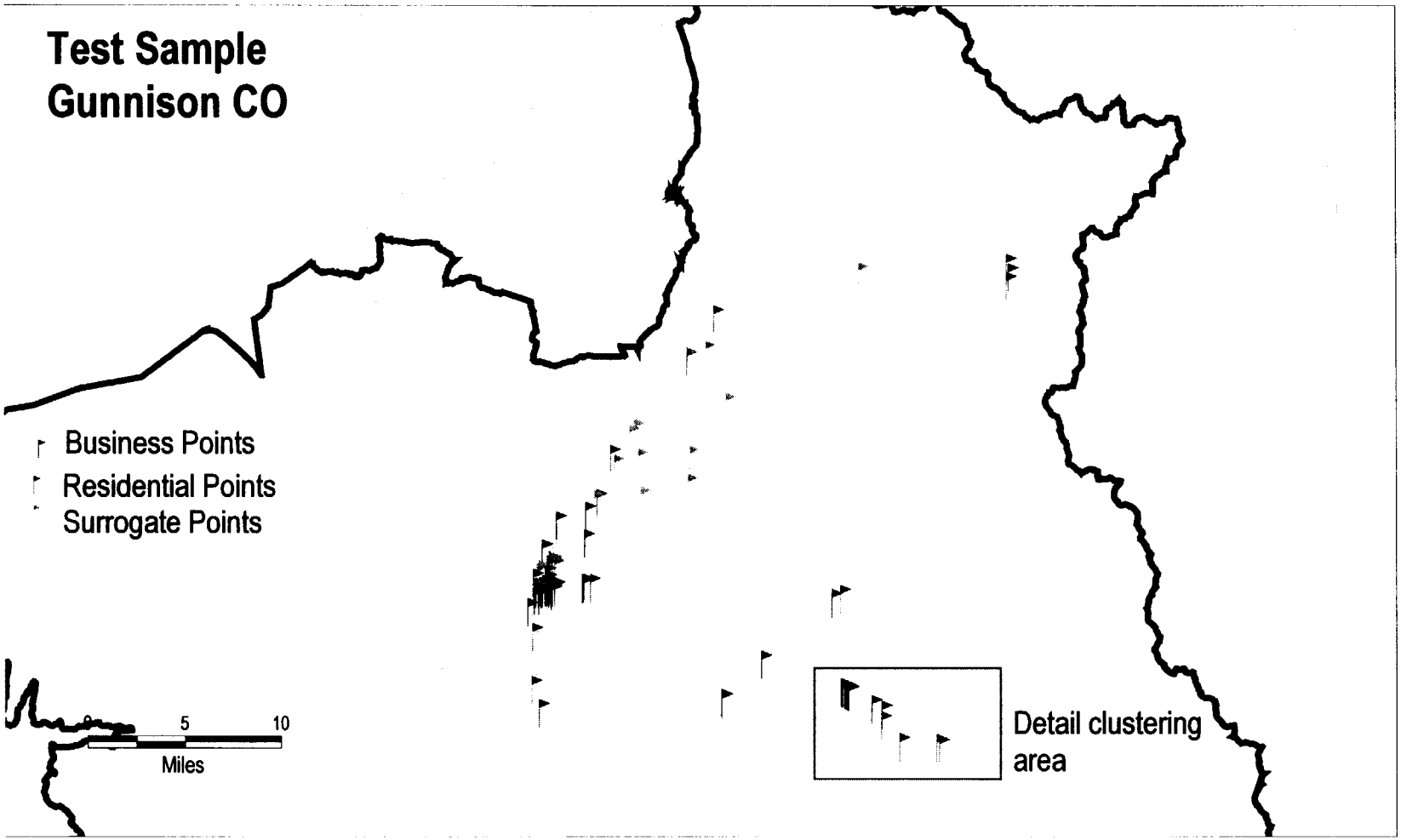
- 100 Residential points
  - 77 records to AS0
  - 9 records to AS1
  - 14 records to a ZIP level
- Leaves customer location file with 86 usable locations

- 100 Business Points
  - 60 records to AS0
  - 8 records to AS1
  - 3 records to AS3
  - 11 records to a ZIP level
  - 18 records could not be geocoded
- Leaves customer location file with 71 usable locations

Customer Location Positioning

- Counts of geocoded customer locations in each CB are compared with WALM target count of customer locations for that CB
  - Count of customer locations that are "unlocated" within the CB is determined
  - "Unlocated" customers are then assumed to be placed uniformly along the periphery of the CB and the geocodes implied by these placements are added to the customer location file (CLF)
  - Assures that the CLF contains geocodes for 100% of the target number of customer locations by CB

Customer Location Positioning



Clustering Customer Locations

- CLF contains a list of customer geocodes for 100% of the target location count for each WC
- To determine efficient engineering of telephone plant, all customer locations associated with a WC are analyzed to determine what cluster patterns exist
  - No artificial CB/CBG or "grid" boundaries are permitted to limit the identification of clusters
- Algorithm must ensure that identified cluster configurations comport with telephone plant engineering specifications

PNR Clustering Algorithm

- Algorithm is based on the "nearest neighbor" methodology with added engineering rules
  - All distances measured based on rectilinear (right angle) routing
  - No point in a cluster may be more than XX kft. distant from the cluster's centroid
  - No cluster may exceed XXXX lines in size
  - No point in a cluster may be farther than X miles from its nearest neighbor in the cluster

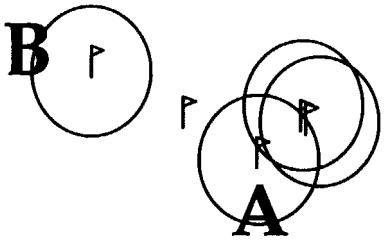
Red parameters are input values to the clustering process

PNR Clustering Algorithm (continued)

- Algorithm allows telephone plant engineering restrictions to be incorporated as an integral part of the clustering process
- Algorithms that require engineering restrictions to be post-applied to clusters are not efficient
  - Efficiency is essential if 111 million customer locations are to be analyzed for clusters

PNR Clustering Algorithm

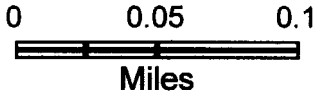
Detail Clustering Area

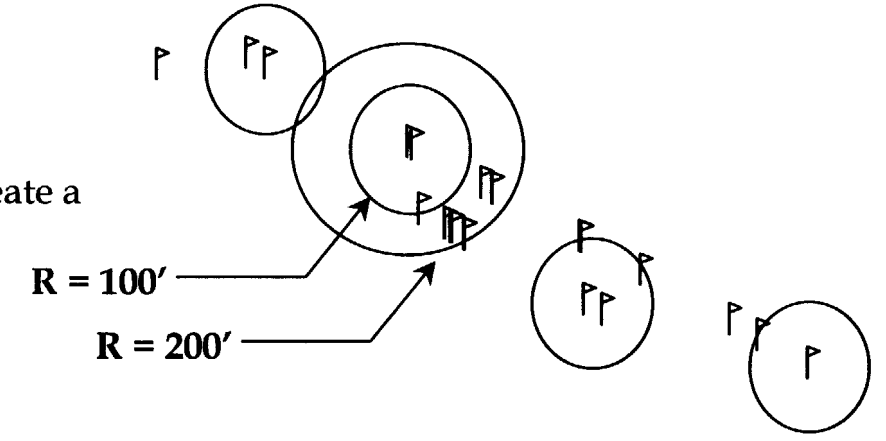


Steps:

1. Start at any point A.
2. Draw a circle 100′ in radius
3. All points within are part of the cluster
4. Continue with each point in the cluster

5. Move to point B and repeat process - complete for all unclustered points.
6. Return to point A and create a circle 200′ in radius.
7. Continue





PNR Clustering Algorithm (continued)

- Engineering requirements for "main" clusters:
  - No point in a main cluster may be more than 18 kft. distant from the main cluster's centroid
  - No main cluster may exceed 1800 lines in size
  - No point in a main cluster may be farther than 1 mile from its nearest neighbor in the main cluster
  - A main cluster must contain at least 5 points
- "Outlier" clusters meet all main cluster requirements, except they contain fewer than 5 points
  - Identity of nearest other cluster established as well

Red parameters are currently used values

PNR Clustering Algorithm